View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

JOURNAL OF MULTIVARIATE ANALYSIS 1, 445–456 (1971)

# Minimum Variance Quadratic Unbiased Estimation of Variance Components

C. RADHAKRISHNA RAO

*Indian Statistical Institute, New Delhi*

The variance of a quadratic function of the random variables in a linear model is minimized to obtain locally best unbiased estimators (MIVQUE) of variance components. Condition for such estimators to be independent of the kurtosis of the variables is given. When the variables are normally distributed, MIVQUE coincides with MINQUE under the Euclidean norm of a matrix. Conditions under which MIVQUE has uniformly minimum variance property are obtained. Expressions are also given for MIMSQE (minimum mean square quadratic estimators).

## 1. INTRODUCTION

In three previous papers [8–10], the author developed a method of estimating variance and covariance components, called MINQUE (Minimum Norm Quadratic Unbiased Estimation). The method is quite general and is applicable in all experimental situations.

In the MINQUE theory a suitable norm of a quadratic form is minimized. When the hypothetical variables in the linear model are scaled by suitable constants as in (7.2), the Euclidean norm corresponds to the variance of the estimator under normal distributions for the variables. In such a case the MINQUEs provide locally minimum variance unbiased quadratic estimators. Harville [4] obtained such estimates in a special case of an unbalanced experiment while solutions were known for balanced experiments [2, 3]. The papers [8–10] provide solutions for general situations. However, the MINQUE theory is developed without reference to normality or variance of the estimator and is highly flexible in the choice of norms, etc.

In the present paper the following problems are considered.

Received June 29, 1971; revised September 2, 1971.

AMS 1970 subject classification: Primary 62J10; Secondary 62H99.

Key words and phrases: Estimation; variance components; minimum variance; minimum mean square unbiased.

(i) Without assuming normality, the variance of an unbiased quadratic form satisfying an invariance condition is minimized at a specified set of values of unknown variance component and kurtosis parameters. Such an estimator is designated as MIVQUE (Minimum Variance Quadratic Unbiased Estimator). A necessary and sufficient condition under which a MIVQUE is independent of unknown kurtosis parameters is obtained.

(ii) Under normality assumption, a necessary and sufficient condition for MINQUE to be a uniformly MIVQUE is obtained.

(iii) In the investigation of (i) and (ii), invariance with respect to location parameter in the linear model is used as an additional condition. MINQUE (or MIVQUE under normality assumption) is obtained without using invariance and under a slightly different norm than that considered in [10].

(iv) Minimum mean square quadratic estimators (MIMSQE) of variance components are obtained in the general case.

The approach of the present paper is a generalization of that of Hsu [5], Rao [6], and the recent work of Drygas [1] on the estimation of the single parameter $\sigma^2$ in the usual Gauss-Markoff model with independent and homoscedastic errors. Naturally the algebraic problem is more complicated in the present case.

## 2. Statement of the Problem

We consider the linear model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}_1\boldsymbol{\xi}_1 + \cdots + \mathbf{U}_k\boldsymbol{\xi}_k\,, \tag{2.1}$$

where the matrices $\mathbf{X}$ of order $n \times m$ and $\mathbf{U}_i$ of order $n \times c_i$, $i = 1,\ldots, k$, are known, $\boldsymbol{\beta}$ is an unknown $m$-vector parameter and $\boldsymbol{\xi}_i$ are hypothetical vector random variables such that

(i) $\boldsymbol{\xi}_i$ is $c_i$-vector with independent components having a common variance $\sigma_i^2$ and kurtosis $\gamma_i$, and

(ii) $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$ are independent. (2.2)

Under the assumptions (2.2), the dispersion matrix of $\mathbf{Y}$ is

$$D(\mathbf{Y}) = \sigma_1^2\mathbf{V}_1 + \cdots + \sigma_k^2\mathbf{V}_k\,, \tag{2.3}$$

where $\mathbf{V}_i = \mathbf{U}_i\mathbf{U}_i'$. The problem we consider is the estimation of a linear function

$$p_1\sigma_1^2 + \cdots + p_k\sigma_k^2 \tag{2.4}$$

of the variance components by a quadratic function $\mathbf{Y'AY}$ satisfying the conditions

$$\begin{aligned} &\text{(i)} \quad \mathbf{AX} = 0, \\ &\text{(ii)} \quad E(\mathbf{Y'AY}) \equiv p_1\sigma_1^2 + \cdots + p_k\sigma_k^2, \qquad (2.5) \\ &\text{(iii)} \quad V(\mathbf{Y'AY}) \text{ is a minimum for a particular choice of } \sigma_1, \dots, \sigma_k. \end{aligned}$$

The motivation for the condition $\mathbf{AX} = 0$ is given in Rao [8–10].

It may be seen that the model (2.1) can be written in a compact form

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}\boldsymbol{\xi}, \tag{2.6}$$

where $\mathbf{U}$ is the $n \times c$, $(c = c_1 + \cdots + c_k)$, partitioned matrix $\mathbf{U} = (\mathbf{U}_1 \vdots \cdots \vdots \mathbf{U}_k)$, and $\boldsymbol{\xi}$ is the $c$-vector $\boldsymbol{\xi}' = (\boldsymbol{\xi}_1' \vdots \cdots \vdots \boldsymbol{\xi}_k')$.

The following expressions for expectation and variance of $\mathbf{Y'AY}$ are easily established:

$$\begin{aligned} E(\mathbf{Y'AY}) &= \operatorname{Tr} \mathbf{A}E(\mathbf{YY'}) \\ &= \operatorname{Tr} \mathbf{A}(\mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}' + \sigma_1^2\mathbf{V}_1 + \cdots + \sigma_k^2\mathbf{V}_k) \\ &= \boldsymbol{\beta}'\mathbf{X}'\mathbf{AX}\boldsymbol{\beta} + \Sigma\sigma_i^2 \operatorname{Tr} \mathbf{AV}_i, \end{aligned} \tag{2.7}$$

where $\mathbf{V}_i = \mathbf{U}_i\mathbf{U}_i'$. The expression for $V(\mathbf{Y'AY})$ when $\mathbf{AX} = 0$

$$V(\mathbf{Y'AY}) = 2 \operatorname{Tr} \mathbf{B}\, \boldsymbol{\Delta}_1 \mathbf{B}\, \boldsymbol{\Delta}_1 + \operatorname{Tr} \tilde{\mathbf{B}}\, \boldsymbol{\Delta}_2 \tilde{\mathbf{B}}, \tag{2.8}$$

where $\mathbf{B} = \mathbf{U'AU}$, $\tilde{\mathbf{B}}$ is the diagonal matrix with the same diagonal elements as in $\mathbf{B}$ and $\boldsymbol{\Delta}_1$, $\boldsymbol{\Delta}_2$ are the diagonal matrices

$$\boldsymbol{\Delta}_1 = \begin{pmatrix} \sigma_1^2\mathbf{I}_1 & & \\ & \ddots & \\ & & \sigma_k^2\mathbf{I}_k \end{pmatrix} \tag{2.9}$$

$$\boldsymbol{\Delta}_2 = \begin{pmatrix} \sigma_1^4\gamma_1\mathbf{I}_1 & & \\ & \ddots & \\ & & \sigma_k^4\gamma_k\mathbf{I}_k \end{pmatrix} \tag{2.10}$$

where $\gamma_i$, as defined earlier, is the kurtosis of the variables in $\boldsymbol{\xi}_i$ and $\mathbf{I}_j$ is the identity matrix of order $c_j$.

If $\mathbf{AX} = \mathbf{0}$, then $E(\mathbf{Y'AY}) = p_1\sigma_1^2 + \cdots + p_k\sigma_k^2$ implies

$$\Sigma\sigma_i^2 \operatorname{Tr} \mathbf{AV}_i = \Sigma p_i\sigma_i^2 \tag{2.11}$$

$$\Rightarrow p_i = \operatorname{Tr} \mathbf{AV}_i, \qquad i = 1, \dots, k. \tag{2.12}$$

Then the problem of MIVQUE reduces to that of minimising

$$2\,\mathrm{Tr}\,\mathbf{B}\,\Delta_1\mathbf{B}\,\Delta_1 + \mathrm{Tr}\,\tilde{\mathbf{B}}\,\Delta_2\tilde{\mathbf{B}} \tag{2.13}$$

subject to the conditions (2.12) and $\mathbf{AX} = \mathbf{0}$.

Let us observe that in some cases it may be possible to find a matrix $\mathbf{Q}$ of maximum rank such that $\mathbf{QX} = \mathbf{0}$, in which case we can eliminate $\mathbf{X}$ from the model (2.1) or (2.6), and obtain the reduced model

$$\mathbf{QY} = \mathbf{QU}\boldsymbol{\xi} \qquad \text{or} \qquad \mathbf{Z} = \mathbf{F}\boldsymbol{\xi}. \tag{2.14}$$

Then the problem is one of finding $\mathbf{Z}'\mathbf{RZ}$ such that

$$E(\mathbf{Z}'\mathbf{RZ}) = p_1\sigma_1^2 + \cdots + p_k\sigma_k^2$$
$$V(\mathbf{Z}'\mathbf{RZ}) \text{ is a minimum.} \tag{2.15}$$

The formulas for such a case are obtained by putting $\mathbf{X} = \mathbf{0}$ and $\mathbf{U} = \mathbf{F}$ in the expressions derived for the general case in Section 3.

We need the following lemma established in Rao [10] which plays an important role in minimising the expression (2.13).

LEMMA 1. *Given a $n \times m$ matrix $\mathbf{X}$ of rank $r$ and a p.d. matrix $\mathbf{V}_*$ of order $n$, there exists a $n \times n - r$ matrix $\mathbf{G}$ of rank $(n - r)$ such that*

$$\mathbf{G}'\mathbf{X} = \mathbf{0} \qquad \textit{and} \qquad \mathbf{G}'\mathbf{V}_*\mathbf{G} = \mathbf{I}. \tag{2.16}$$

*For any matrix $\mathbf{G}$ satisfying* (2.16)

$$\mathbf{GG}' = \mathbf{V}_*^{-1}(\mathbf{I} - \mathbf{P}_v), \tag{2.17}$$

*where*

$$\mathbf{P}_v = \mathbf{X}(\mathbf{X}'\mathbf{V}_*^{-1}\mathbf{X})^-\,\mathbf{X}'\mathbf{V}_*^{-1} \tag{2.18}$$

*is the projection operator onto the space generated by the columns of $\mathbf{X}$.*

*Notation.* In solving the minimisation problem in Section 3, the matrix defined in (2.16) is brought in only as an intermediate step. The final results are expressible in terms of $\mathbf{GG}'$ only, which involves the projection operator $\mathbf{P}_v$ as in (2.17). We shall give the notations used in Sections 3 and 4 for ready reference.

For any matrix $\mathbf{B}$, we denote by $\tilde{\mathbf{B}}$ the diagonal matrix with the same diagonal elements as in $\mathbf{B}$. We denote

$$\mathbf{M} = \mathbf{GG}'\mathbf{U} = \mathbf{V}_*^{-1}(\mathbf{I} - \mathbf{P}_v)\mathbf{U}, \tag{2.19}$$
$$\mathbf{M}_1 = \mathbf{U}'\mathbf{M} = (m_{ij}), \tag{2.20}$$
$$\mathbf{M}_2 = (m_{ij}^2). \tag{2.21}$$

We choose $\mathbf{V}_* = \mathbf{U}\,\Delta_1\mathbf{U}' = \Sigma\sigma_i^2\mathbf{V}_i$ in defining $\mathbf{G}$.

## 3. Algebraic Derivation of MIVQUE

The variance (2.13) of $\mathbf{Y'AY}$ will be a minimum if the matrix $\mathbf{A}$ is such that

$$\operatorname{Cov}(\mathbf{Y'AY}, \mathbf{Y'NY}) = 0 \tag{3.1}$$

for any $\mathbf{N}$ such that $E(\mathbf{Y'NY}) = 0$. The proof of the statement can be found in Rao [6, p. 28; 7, p. 257]. Let $\mathbf{U'AU} = \mathbf{B}$ and $\mathbf{U'NU} = \mathbf{F}$. Then assuming $\mathbf{AX} = 0$ and $\mathbf{NX} = 0$,

$$\operatorname{Cov}(\mathbf{Y'AY}, \mathbf{Y'NY}) = 2\operatorname{Tr} \mathbf{B}\, \mathbf{\Delta}_1 \mathbf{F}\, \mathbf{\Delta}_1 + \operatorname{Tr} \tilde{\mathbf{B}}\, \mathbf{\Delta}_2 \tilde{\mathbf{F}}, \tag{3.2}$$

where $\mathbf{\Delta}_1$ and $\mathbf{\Delta}_2$ are as defined in (2.9) and (2.10), respectively. Under the conditions $\mathbf{AX} = \mathbf{0}$ and $\mathbf{NX} = \mathbf{0}$, $\mathbf{A}$ can be written as $\mathbf{GCG'}$ and $\mathbf{N}$ as $\mathbf{GDG'}$, where $\mathbf{G}$ may be chosen to satisfy the conditions

$$\mathbf{G'X} = \mathbf{0}, \qquad \mathbf{G'U}\, \mathbf{\Delta}_1 \mathbf{U'G} = \mathbf{I} \tag{3.3}$$

as in Lemma 1. Denoting the $i$-th column of $\mathbf{G'U}$ by $\boldsymbol{\alpha}_i$ and the $i$-th diagonal element of $\mathbf{\Delta}_2$ by $\delta_i$ the expression (3.2) can be written as

$$\operatorname{Tr} \mathbf{D}(2\mathbf{C} + \Sigma\, \delta_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i' \mathbf{C} \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i'). \tag{3.4}$$

Now

$$E(\mathbf{Y'GDG'Y}) \equiv 0 \Rightarrow \operatorname{Tr} \mathbf{D}(\Sigma \nu_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i') = 0 \tag{3.5}$$

for all $\nu_1, \nu_2, \ldots$ of the form

$$\begin{pmatrix} \nu_1 & & \\ & \ddots & \\ & & \nu_c \end{pmatrix} = \begin{pmatrix} \lambda_1 \mathbf{I}_1 & & \\ & \ddots & \\ & & \lambda_k \mathbf{I}_k \end{pmatrix} = \mathbf{\Lambda}, \tag{3.6}$$

where $\lambda_1, \ldots, \lambda_k$ are arbitrary. If the expression (3.4) has to be zero for all $\mathbf{D}$ satisfying (3.5), then we must have

$$2\mathbf{C} + \Sigma\, \delta_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i' \mathbf{C} \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i' = \Sigma \nu_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i' \tag{3.7}$$

which is the fundamental equation providing the optimum matrix $\mathbf{C}$. Writing $\theta_i = \boldsymbol{\alpha}_i' \mathbf{C} \boldsymbol{\alpha}_i$, Eq. (3.7) can be written

$$2\mathbf{C} + \Sigma\, \delta_i \theta_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i' = \Sigma \nu_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i'. \tag{3.8}$$

Multiplying both sides of (3.8) by $\boldsymbol{\alpha}_j'$ from the left and $\boldsymbol{\alpha}_j$ from the right we obtain

$$2\theta_j + \Sigma\, \delta_i \theta_i m_{ij}^2 = \Sigma \nu_i m_{ij}^2, \qquad j = 1, \ldots, c, \tag{3.9}$$

where $m_{ij} = \boldsymbol{\alpha}_i'\boldsymbol{\alpha}_j$ as in (2.20). In matrix notation, (3.9) can be written as

$$(2\mathbf{I} + \mathbf{M}_2\,\boldsymbol{\Delta}_2)\boldsymbol{\theta} = \mathbf{M}_2\boldsymbol{\nu}, \tag{3.10}$$

where $\mathbf{M}_2 = (m_{ij}^2)$, $\boldsymbol{\theta}' = (\theta_1, \ldots, \theta_c)$, $\nu' = (\nu_1, \ldots, \nu_c)$. The condition of unbiasedness

$$E(\mathbf{Y}'\mathbf{GCG}'\mathbf{Y}) \equiv p_1\sigma_1^2 + \cdots + p_k\sigma_k^2 \tag{3.11}$$

yields the equation

$$\mathbf{H}\boldsymbol{\theta} = \mathbf{p}, \tag{3.12}$$

where $\mathbf{p}' = (p_1, \ldots, p_k)$ and the $i$-th row of $\mathbf{H}$ has unity in the $c_i$ positions beginning from the $(c_1 + \cdots + c_{i-1} + 1)$th and zeroes elsewhere. Let us observe that $\nu$ has only $k$ distinct components $\lambda_1, \ldots, \lambda_k$ so that $\mathbf{M}_2\boldsymbol{\nu} = \mathbf{W}\boldsymbol{\lambda}$ where $\mathbf{W}$ is obtained by appropriate summing up of the columns of $\mathbf{M}_2$. Then we have the two equations

$$\begin{aligned} (\mathbf{M}_2\,\boldsymbol{\Delta}_2 + 2\mathbf{I})\boldsymbol{\theta} - \mathbf{W}\boldsymbol{\lambda} &= \mathbf{0} \\ \mathbf{H}\boldsymbol{\theta} &= \mathbf{p} \end{aligned} \tag{3.13}$$

providing the values of $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$ which specify $\mathbf{C}$ through Eq. (3.8), and thus the quadratic form $\mathbf{Y}'\mathbf{GCG}'\mathbf{Y}$ which is the MIVQUE of $p_1\sigma_1^2 + \cdots + p_k\sigma_k^2$. The result is summarized in the following theorem where $\boldsymbol{\Lambda}$ is as defined in (3.6) and $\boldsymbol{\Theta}$ is a diagonal matrix with the components of vector $\boldsymbol{\theta}$ as its diagonal elements.

THEOREM 1. *The MIVQUE of $\Sigma p_i\sigma_i^2$ is*

$$\mathbf{Y}'\mathbf{M}(\tfrac{1}{2}\boldsymbol{\Lambda} - \tfrac{1}{2}\,\boldsymbol{\Delta}_2\boldsymbol{\Theta})\mathbf{M}'\mathbf{Y}, \tag{3.14}$$

*where* $\mathbf{M}$ *is as defined in* (2.19). *The vectors* $\boldsymbol{\lambda}$ *and* $\boldsymbol{\theta}$ *which determine* $\boldsymbol{\Lambda}$ *and* $\boldsymbol{\Theta}$ *satisfy the equation*

$$\begin{aligned} (2\mathbf{I} + \mathbf{M}_2\,\boldsymbol{\Delta}_2)\boldsymbol{\theta} - \mathbf{W}\boldsymbol{\lambda} &= \mathbf{0} \\ \mathbf{H}\boldsymbol{\theta} &= \mathbf{p} \end{aligned} \tag{3.15}$$

*where* $\mathbf{M}_2$ *is as defined in* (2.21).

Equation (3.8) can be written in the matrix form

$$2\mathbf{C} + \mathbf{G}'\mathbf{U}\,\boldsymbol{\Delta}_2\boldsymbol{\Theta}\mathbf{U}'\mathbf{G} = \mathbf{G}'\mathbf{U}\,\boldsymbol{\Lambda}\mathbf{U}'\mathbf{G}. \tag{3.16}$$

Multiplying Eq. (3.16) from left and right by $\mathbf{G}$ and $\mathbf{G}'$, respectively, we have

$$2\mathbf{GCG}' + \mathbf{M}\,\boldsymbol{\Delta}_2\boldsymbol{\Theta}\mathbf{M}' = \mathbf{M}\,\boldsymbol{\Lambda}\mathbf{M}', \tag{3.17}$$

since $\mathbf{M} = \mathbf{GG'U}$. From (3.17)

$$2\mathbf{GCG'} = \mathbf{M}(\mathbf{\Lambda} - \mathbf{\Delta}_2\mathbf{\Theta})\mathbf{M'} \tag{3.18}$$

which proves the desired result.

THEOREM 2. *The MIMSQE of* $\Sigma p_i\sigma_i^2$ *is*

$$(\tfrac{1}{2})\,\mathbf{Y'}[d\mathbf{V}_*^{-1}(1 - \mathbf{P}_v) - \mathbf{M\Delta}_2\mathbf{\Theta M'}]\mathbf{Y}$$

*where* $\boldsymbol{\theta}$ *which determines* $\mathbf{\Theta}$ *and the constant* $d$ *satisfy the equations*

$$\begin{aligned}(2\mathbf{I} + \mathbf{M}_2\mathbf{\Delta}_2)\,\boldsymbol{\theta} - d\tilde{\mathbf{M}}_1 &= 0,\\ \operatorname{Tr} \mathbf{\Delta}_1\boldsymbol{\theta} + d &= \Sigma p_i\sigma_i^2.\end{aligned} \tag{3.19}$$

When the variables are normally distributed, the MIMSQE of $\Sigma p_i\sigma_i^2$ is

$$\frac{\Sigma p_i\sigma_i^2}{n - r + 2}\,\mathbf{Y'}[\mathbf{V}_*^{-1}(1 - \mathbf{P}_v)]\mathbf{Y} \tag{3.20}$$

where $r$ is the rank of $X_v$ and $\mathbf{V}_1 = \Sigma\mathbf{V}_i\sigma_i^2$.

## 4. SIMULTANEOUS ESTIMATION OF VARIANCE COMPONENTS

Theorem 1 as stated in Section 3 refers to the estimation of a particular parametric function and is not in line with previous work where estimates of $\sigma_1^2,\ldots, \sigma_k^2$ are obtained first and then substituted in a given parametric function. We shall provide a similar technique in the present case also. From (3.8),

$$2\mathbf{C} = -\Sigma\,\delta_i\theta_i\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' + \nu_i\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' \tag{4.1}$$

we obtain the MIVQUE in the form

$$2\mathbf{Y'GCG'Y} = \Sigma(\nu_i - \delta_i\theta_i)u_i^2 \tag{4.2}$$

where $u_i = \boldsymbol{\alpha}_i'\mathbf{GY}$, the $i$-th component of $\mathbf{U'GG'Y} = \mathbf{M'Y}$. Let us denote the vectors

$$\begin{aligned}\mathbf{w'} &= (\delta_1u_1^2,\ldots, \delta_cu_c^2),\\ \mathbf{t'} &= (\Sigma_1u_i^2,\ldots, \Sigma_ku_i^2),\end{aligned} \tag{4.3}$$

where $\Sigma_1$ represents summation over $i = 1$ to $c_1$, $\Sigma_2$ over $i = c_1 + 1$ to $c_2$, etc. Then from (4.2), the estimate is of the form

$$\tfrac{1}{2}(-\boldsymbol{\theta}'\mathbf{w} + \boldsymbol{\lambda}'\mathbf{t}), \tag{4.4}$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$ are a solution of

$$\begin{aligned}(2\mathbf{I} + \mathbf{M}_2\, \boldsymbol{\Delta}_2)\boldsymbol{\theta} - \mathbf{W}\boldsymbol{\lambda} &= 0 \\ \mathbf{H}\boldsymbol{\theta} &= \mathbf{p}.\end{aligned} \tag{4.5}$$

If

$$\begin{pmatrix} \mathbf{E}_1 & \mathbf{E}_2 \\ \mathbf{E}_3 & \mathbf{E}_4 \end{pmatrix} \tag{4.6}$$

is a $g$-inverse (see Rao [7]) of the matrix of equations in (4.5) then

$$\hat{\boldsymbol{\theta}} = \mathbf{E}_2\mathbf{p}, \qquad \hat{\boldsymbol{\lambda}} = \mathbf{E}_4\mathbf{p} \tag{4.7}$$

is a solution. Substituting (4.7) in (4.4), the estimate of $\Sigma p_i\sigma_i^2$ is

$$\tfrac{1}{2}p'(-\mathbf{E}_2'\mathbf{w} + \mathbf{E}_4'\mathbf{t}) \tag{4.8}$$

which shows that the individual estimates of $\sigma_i^2$ (if estimable) are the components of $(\frac{1}{2})(-\mathbf{E}_2'\mathbf{w} + \mathbf{E}_4'\mathbf{t})$.

*Note.* In all the formulas for estimating $\Sigma p_i\sigma_i^2$ the true values of $\sigma_i^2$ appear. In practice we use a priori values or a given set of values at which a minimum is sought.

## 5. When is MINQUE a MIVQUE?

In MINQUE, the expression whose minimum is sought is

$$\operatorname{Tr} \mathbf{B}\, \boldsymbol{\Delta}_1 \mathbf{B}\, \boldsymbol{\Delta}_1 \tag{5.1}$$

while in MIVQUE the corresponding expression is

$$2\operatorname{Tr} \mathbf{B}\, \boldsymbol{\Delta}_1 \mathbf{B}\, \boldsymbol{\Delta}_1 + \operatorname{Tr} \tilde{\mathbf{B}}\, \boldsymbol{\Delta}_2 \tilde{\mathbf{B}}, \tag{5.2}$$

where $\mathbf{B}$, $\boldsymbol{\Delta}_1$, $\boldsymbol{\Delta}_2$ are as defined in (2.13). To answer the problem raised we have to determine the conditions under which (5.1) and (5.2) attain a minimum for the same matrix $\mathbf{B}$ or, in other words, the conditions under which the minimising matrix of (5.2) is independent of $\boldsymbol{\Delta}_2$ or the kurtosis of the structural variables $\boldsymbol{\xi}$.

From (3.8) we have

$$2\mathbf{C} + \Sigma\, \delta_i\theta_i\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' = \Sigma\nu_i\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i'. \tag{5.3}$$

If $\delta_i = 0$ for all $i$, then

$$2\mathbf{C} = \Sigma\eta_i\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' \tag{5.4}$$

for a suitable choice of $\eta_i$. Substituting the expression (5.4) for $\mathbf{C}$ in (5.3),

$$\Sigma\eta_i\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' + \Sigma\,\delta_i\theta_i\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' = \Sigma\nu_i\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i' \tag{5.5}$$

which implies, by the process used in deriving (3.10) from (3.8),

$$\mathbf{W}\boldsymbol{\mu} + \mathbf{M}_2\,\boldsymbol{\Delta}_2\boldsymbol{\theta} = \mathbf{W}\boldsymbol{\lambda}, \tag{5.6}$$

where $\boldsymbol{\mu}$ is the vector of distinct $\eta_i$ and the matrices $\mathbf{M}_2$ and $\mathbf{W}$ are as in (4.5). Using (4.5) and (5.6) we find

$$2\boldsymbol{\theta} = \mathbf{W}\boldsymbol{\mu}. \tag{5.7}$$

Substituting for $\boldsymbol{\theta}$ in (5.6),

$$\mathbf{M}_2\,\boldsymbol{\Delta}_2\mathbf{W}\boldsymbol{\mu} = 2\mathbf{W}(\boldsymbol{\lambda} - \boldsymbol{\mu}) = \mathbf{W}\boldsymbol{\nu}. \tag{5.8}$$

Partitioning

$$\mathbf{M}_2 = (\mathbf{W}_1 \,\vdots\, \cdots \,\vdots\, \mathbf{W}_k), \tag{5.9}$$

where $\mathbf{W}_i$ is a matrix of order $c \times c_i$ and

$$\mathbf{W}' = (\mathbf{N}_1' \,\vdots\, \cdots \,\vdots\, \mathbf{N}_k'), \tag{5.10}$$

Eq. (5.8) can be written

$$(\zeta_1\mathbf{W}_1\mathbf{N}_1 + \cdots + \zeta_k\mathbf{W}_k\mathbf{N}_k)\boldsymbol{\mu} = \mathbf{W}\boldsymbol{\nu} \tag{5.11}$$

where $\zeta_1, \ldots, \zeta_k$ are the distinct diagonal elements of $\boldsymbol{\Delta}_2$. If Eq. (5.11) holds for all $\zeta_i$ and $\boldsymbol{\mu}$, the space generated by the columns of $\mathbf{W}_i\mathbf{N}_i$ for each $i$ must belong to the space generated by the columns of $\mathbf{W}$, which is the necessary and sufficient condition that the MINQUE of any estimatable linear function of the variance components is also the MIVQUE.

## 6. The Case of Normal Variables (with Invariance)

When the hypothetical variables $\boldsymbol{\xi}_i$ in (2.1) are normally distributed the expression for $V(\mathbf{Y}'\mathbf{A}\mathbf{Y})$ under the condition $\mathbf{AX} = \mathbf{0}$ reduces to

$$2\operatorname{Tr}\mathbf{A}\mathbf{V}_*\mathbf{A}\mathbf{V}_* \tag{6.1}$$

where $\mathbf{V}_* = \sigma_1^2\mathbf{V}_1 + \cdots + \sigma_k^2\mathbf{V}_k$. In such a case the MIVQUE and the MINQUE estimators are the same. It would be of some interest to find conditions under which the matrix $\mathbf{A}$ minimizing (6.1) is independent of the unknown parameters $\sigma_1^2, \ldots, \sigma_k^2$, so that uniformly minimum variance estimators are available. Then MINQUE is unique whatever $\sigma_i^2$ may be in (6.1).

Let $\mathbf{V} = \mathbf{V}_1 + \cdots + \mathbf{V}_k$. It is shown in [10] that min Tr $\mathbf{AVAV}$ subject to $\mathbf{AX} = \mathbf{0}$ and Tr $\mathbf{AV}_i = p_i$, $i = 1,..., k$ is attained at

$$\mathbf{A}_* = \Sigma\lambda_j \mathbf{Q}_v'\mathbf{V}^{-1}\mathbf{V}_j\mathbf{V}^{-1}\mathbf{Q}_v, \tag{6.2}$$

where $\mathbf{Q}_v = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}$ and $\lambda_j$ are chosen to satisfy the conditions Tr $\mathbf{AV}_i = p_i$, $i = 1,..., k$. The expression Tr $\mathbf{AV}_*\mathbf{AV}_*$ will be a minimum at $\mathbf{A}_*$ determined in (6.2) if $\mathbf{B}_* = \mathbf{Q}_v'\mathbf{V}^{-1}\mathbf{V}_j\mathbf{V}^{-1}\mathbf{Q}_v$ minimizes Tr $\mathbf{BV}_*\mathbf{BV}_*$ subject to $\mathbf{BX} = \mathbf{0}$ and Tr $\mathbf{BV}_i$ = Tr $\mathbf{B}_*\mathbf{V}_i$, $i = 1,..., k$.

Let $\mathbf{D}$ be a symmetric matrix such that $\mathbf{DX} = \mathbf{0}$ and Tr $\mathbf{DV}_i = 0$, $i = 1,..., k$. The condition $\mathbf{DX} = 0 \Leftrightarrow \mathbf{D} = \mathbf{Q}_v'\mathbf{EQ}_v$ and Tr $\mathbf{DV}_i = 0 \Leftrightarrow$ Tr $\mathbf{EQ}_v\mathbf{V}_i\mathbf{Q}_v' = 0$. If $\mathbf{B}_*$ minimizes Tr $\mathbf{BV}_*\mathbf{BV}_*$, then it is necessary and sufficient that

$$\begin{aligned} \operatorname{Tr} \mathbf{B}_*\mathbf{V}_*\mathbf{DV}_* &= \operatorname{Tr} \mathbf{Q}_v'\mathbf{V}^{-1}\mathbf{V}_j\mathbf{V}^{-1}\mathbf{Q}_v\mathbf{V}_*\mathbf{Q}_v'\mathbf{EQ}_v\mathbf{V}_* \\ &= \operatorname{Tr} \mathbf{E}(\mathbf{Q}_v\mathbf{V}_*\mathbf{Q}_v'\mathbf{V}^{-1}\mathbf{V}_j\mathbf{V}^{-1}\mathbf{Q}_v\mathbf{V}_*\mathbf{Q}_v') = 0, \end{aligned} \tag{6.3}$$

whenever Tr $\mathbf{EQ}_v\mathbf{V}_i\mathbf{Q}_v' = 0$, $i = 1,..., k$. Then it follows that

$$\mathbf{Q}_v\mathbf{V}_*\mathbf{Q}_v'\mathbf{V}^{-1}\mathbf{V}_j\mathbf{V}^{-1}\mathbf{Q}_v\mathbf{V}_*\mathbf{Q}_v' = \Sigma\mu_i\mathbf{Q}_v\mathbf{V}_i\mathbf{Q}_v' \tag{6.4}$$

for a suitable choice of constants $\mu_1,..., \mu_k$. The relation (6.4) must hold for all values of $\sigma_1^2,..., \sigma_k^2$ which define $\mathbf{V}_*$, which implies that

$$\mathbf{Q}_v\mathbf{V}_i\mathbf{Q}_v'\mathbf{V}^{-1}\mathbf{V}_j\mathbf{V}^{-1}\mathbf{Q}_v\mathbf{V}_s\mathbf{Q}_v' \tag{6.5}$$

is a linear combination of $\mathbf{Q}_v\mathbf{V}_r\mathbf{Q}_v'$, $r = 1,..., k$, for all triplets $i, j, s$. No further simplification of the condition (6.5) seems to be possible.

If we work with the model (2.15) after eliminating $\mathbf{X\beta}$ then the condition is that

$$\mathbf{V}_i\mathbf{V}^{-1}\mathbf{V}_j\mathbf{V}^{-1}\mathbf{V}_s \tag{6.6}$$

is a linear combination of $\mathbf{V}_r$ for all triplets $i, j, s$.

## 7. The Case of Normal Variables (without Invariance)

Let us consider the model (2.1)

$$\mathbf{Y} = \mathbf{X\beta} + \mathbf{U}_1\boldsymbol{\xi}_1 + \cdots + \mathbf{U}_k\boldsymbol{\xi}_k. \tag{7.1}$$

By suitable scaling of $\boldsymbol{\xi}_i$ and transformation of $\boldsymbol{\beta}$ we rewrite (7.1) as

$$\mathbf{Y} - \mathbf{X\beta}_0 = \mathbf{XF\gamma} + \alpha_1\mathbf{U}_1\boldsymbol{\eta}_1 + \cdots + \alpha_k\mathbf{U}_k\boldsymbol{\eta}_k \tag{7.2}$$

$$= \mathbf{XF\gamma} + \mathbf{U}_*\boldsymbol{\eta},\ \mathbf{U}_* = (\alpha_1\mathbf{U}_1 \vdots \cdots \vdots \alpha_k\mathbf{U}_k) \tag{7.3}$$

such that

(i) $\boldsymbol{\eta}_i$ has $I$ as dispersion matrix when $\alpha_i = \sigma_i$;

(ii) $\boldsymbol{\gamma}$ has $I$ as dispersion matrix when $E(\boldsymbol{\beta}) = \boldsymbol{\beta}_0$ and $D(\boldsymbol{\beta}) = \mathbf{FF}'$, considering $\boldsymbol{\beta}$ as a random variable or $\mathbf{F}\boldsymbol{\gamma} = \boldsymbol{\beta} - \boldsymbol{\beta}_0$ when $\boldsymbol{\beta}$ is considered as a vector of constants. The matrices and constants $\mathbf{F}$, $\boldsymbol{\beta}_0$, $\alpha_1, \ldots, \alpha_k$ constitute our *a priori* knowledge about the unknown quantities $\boldsymbol{\beta}$, $\sigma_1, \ldots, \sigma_k$ and are considered as imputs of the problem. In the absence of any *a priori* knowledge we may have to work with (7.1) instead of (7.2).

We consider the class $\mathscr{C}_4$ of quadratic estimators $(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0)'\mathbf{A}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0)$ where $\mathbf{A}$ satisfies the conditions

$$\mathbf{X}'\mathbf{A}\mathbf{X} = 0, \qquad \operatorname{Tr} \mathbf{A}\mathbf{V}_i = p_i, \qquad i = 1, \ldots, k \tag{7.4}$$

as in [10]. The quadratic form $(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0)'\mathbf{A}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0)$ under the condition $\mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{0}$ can be written

$$\boldsymbol{\eta}'\mathbf{U}_*'\mathbf{A}\mathbf{U}_*\boldsymbol{\eta} + 2\boldsymbol{\eta}'\mathbf{U}_*'\mathbf{A}\mathbf{X}\mathbf{F}\boldsymbol{\gamma}. \tag{7.5}$$

The MINQUE theory requires the minimization of norm of the matrix

$$\begin{pmatrix} \mathbf{U}_*'\mathbf{A}\mathbf{U}_* & \mathbf{U}_*'\mathbf{A}\mathbf{X}\mathbf{F} \\ \mathbf{F}'\mathbf{X}'\mathbf{A}\mathbf{U}_* & \cdot \end{pmatrix}. \tag{7.6}$$

Using Euclidean norm the expression is

$$\operatorname{Tr} \mathbf{A}\mathbf{V}_*\mathbf{A}\mathbf{V}_* + 2\operatorname{Tr} \mathbf{A}\mathbf{X}\mathbf{F}\mathbf{F}'\mathbf{X}'\mathbf{A}\mathbf{V}_* = \operatorname{Tr} \mathbf{A}(\mathbf{V}_* + 2\mathbf{X}\mathbf{F}\mathbf{F}'\mathbf{X}')\mathbf{A}\mathbf{V}_*, \tag{7.7}$$

where $\mathbf{V}_* = \mathbf{U}_*\mathbf{U}_*'$. It may be noted that (7.7) is proportional to the variance of $(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0)'\mathbf{A}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0)$ when $\sigma_i = \alpha_i$, $\mathbf{FF}' = (\boldsymbol{\beta} - \boldsymbol{\beta}_0)(\boldsymbol{\beta} - \boldsymbol{\beta}_0)'$ and the hypothetical variables have normal distributions.

The problem of minimizing (7.7) for a general $\mathbf{FF}' = D(\boldsymbol{\beta}) = E(\boldsymbol{\beta} - \boldsymbol{\beta}_0)(\boldsymbol{\beta} - \boldsymbol{\beta}_0)'$ is solved in Lemma 3.12 of [10]. The optimum choice of $\mathbf{A}$ is

$$\mathbf{A}_* = \mathbf{N}(\Sigma\lambda_i\mathbf{W}_i)\mathbf{N}' - \mathbf{P}_v'\mathbf{N}(\Sigma\lambda_i\mathbf{W}_i)\mathbf{N}'\mathbf{P}_v, \tag{7.8}$$

where $\mathbf{N}$ is such that $\mathbf{N}\mathbf{V}_*\mathbf{N}' = \mathbf{I}$, $\mathbf{N}(\mathbf{V}_* + 2\mathbf{X}\mathbf{F}\mathbf{F}'\mathbf{X}')\mathbf{N}' = \boldsymbol{\Delta}$ (diagonal with diagonal elements $\delta_1, \ldots, \delta_n$), $\mathbf{P}_v = \mathbf{X}(\mathbf{X}'\mathbf{V}_*^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{V}_*^{-1}$, and $\mathbf{W}_i$ is a matrix such that its $(r, s)$-th element is the $(r, s)$-th element of $\mathbf{N}'\mathbf{V}_i\mathbf{N}$ divided by $(\delta_r + \delta_s)$. The $\lambda_i$ are chosen subject to the conditions, $\operatorname{Tr} \mathbf{A}_*\mathbf{V}_i = p_i$, $i = 1, \ldots, k$.

In the special case when $\mathbf{FF}' = (\boldsymbol{\beta} - \boldsymbol{\beta}_0)(\boldsymbol{\beta} - \boldsymbol{\beta}_0)'$, i.e., $R(\mathbf{FF}') = 1$, the optimum choice of $\mathbf{A}$ reduces to the simpler form

$$\mathbf{A}_* = \mathbf{V}_*^{-1}(\Sigma\lambda_i\mathbf{B}_i)\mathbf{V}_*^{-1}, \tag{7.9}$$

where

$$\begin{aligned}\mathbf{B}_i &= (\mathbf{V}_i - \mathbf{P}_v\mathbf{V}_i\mathbf{P}_v{}') - (1 + c)^{-1}(\mathbf{V}_i\mathbf{V}_*^{-1}\mathbf{K} - \mathbf{P}_v\mathbf{V}_i\mathbf{V}_*^{-1}\mathbf{K}\mathbf{P}_v{}'),\\ \mathbf{P}_v &= \mathbf{X}(\mathbf{X}'\mathbf{V}_*^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}_*^{-1},\\ c &= (\boldsymbol{\beta} - \boldsymbol{\beta}_0)'\mathbf{X}'\mathbf{X}(\boldsymbol{\beta} - \boldsymbol{\beta}_0),\\ \mathbf{K} &= \mathbf{X}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)(\boldsymbol{\beta} - \boldsymbol{\beta}_0)'\mathbf{X}'.\end{aligned} \tag{7.10}$$

In (7.10), special choices of $\beta = 0$ and $\beta_0 = 0$ are of interest.

When the hypothetical variables do not have a normal distribution, the variance of the quadratic form involves the third and fourth moments. Minimization in such cases is under investigation.

*Note added in proof.* It has come to the notice of the author that special cases of the results of the paper when then kurtosis of the variables is zero are obtained by LaMotte [11]. But many of these special cases are already covered in [10] by the author.

## References


[1] Drygas, H. (1970). The estimation of residual variance in regression analysis. CORE Discussion paper No. 7011.

[2] Graybill, F. A. (1954). On quadratic estimates of variance components. *Ann. Math. Statist.* **25** 367–372.

[3] Graybill, F. A. and Hulkquist, R. A. (1961). Theorems concerning Eisenharts' model II. *Ann. Math. Statist.* **32** 261–269.

[4] Harville, David A. (1969). Quadratic unbiased estimation of variance components for the one-way classification. *Biometrika* **56** 313–326.

[5] Hsu, P. L. (1938). On the best unbiased quadratic estimate of variance. *Statist. Res. Mem.* **2** 91–104.

[6] Rao, C. Radhakrishna (1952). Some theorems on minimum variance estimation. *Sankhyā* **12** 27–56.

[7] Rao, C. Radhakrishna (1965). *Linear Statistical Inference and Its Applications.* Wiley, New York.

[8] Rao, C. Radhakrishna (1970). Estimation of heteroscedastic variances in linear models. *J. Amer. Statist. Assoc.* **65** 161–172.

[9] Rao, C. Radhakrishna (1971a). Estimation of variance and covariance components in linear models. *J. Amer. Statist. Assoc.*, to appear.

[10] Rao, C. Radhakrishna (1971b). Estimation of variance and covariance components—MINQUE theory. *J. Multivariate Anal.* **1** 257–275.

[11] La Motte, L. R. (1971). Locally best quadratic estimators of variance components. Technical report.